Store Expensiveness and Consumer Saving: Insights from a New Decomposition of Price Dispersion

Sofronis Clerides[1,4] Pascal Courty[2,4] Yupei Ma[3]

[1]University of Cyprus
[2]University of Victoria
[3]University of British Columbia
[4]Centre for Economic Policy Research (CEPR)

20th CRETE Conference
11-15 July 2022

Table of contents

1. Introduction
2. Our contribution
3. The method
4. Understanding the decomposition
5. An example
6. Concluding remarks

Price dispersion and consumer saving I

- Price dispersion: when the same good is available for purchase at different prices
- A well-documented phenomenon in many datasets and markets
- In particular, several recent studies using detailed grocery store data find that prices vary considerably
  - across time within a given store
  - across stores within the same time period
- In principle, price-sensitive consumers can exploit price differences to purchase their desired basket of goods at a lower overall cost
- Do they? What kind of price variation do they exploit more effectively? What is the right shopping model?

Price dispersion and consumer saving II

- Important questions; people spend a lot of money in grocery stores
- Old literature on intertemporal price dispersion (sales promotions) $\Rightarrow$ substantial saving opportunity
- More recent literature focuses on the multi-product and multi-store nature of grocery shopping
- *Relative price dispersion*: persistent differences in the price that retailers set for one particular good relative to the price they set for other goods
  (Kaplan and Menzio (2015), Kaplan, Menzio, Rudanko and Trachter (2019))
- Consumers with low search costs can save relative to those who shop from a single store or do not compare prices across stores

How do people save? I

Kaplan and Menzio (2015): *The morphology of price dispersion*

- Main emphasis on documenting price dispersion in grocery stores
- Also analyze variation in prices consumers pay
- Data:
  - All purchases made by large sample of households
  - Prices of all products in (almost) all stores
- KM calculate a household price index (HPI):

$$\text{HPI} = \frac{\text{actual cost of a consumer's basket}}{\text{cost of same basket at average market price of each product}}$$

- $\text{HPI} = 1 \Rightarrow$ consumer paid the average market price for her basket

How do people save? II

- They decompose the HPI into three components:
  - **Store component**: difference between prices at stores visited and overall market prices $\Rightarrow$ **store choice**
  - **Transaction component**: difference between average price paid for item and average price of same item in same store $\Rightarrow$ **purchase timing**
  - **Store-good component**: measure of the expensiveness of products purchased by the consumer relative to the store's overall expensiveness; $\Rightarrow$ **cross-store shopping**

How do people save? III

- Then they compute the variance of the HPI that is due to each component
- Findings:
  - **50%** due to store-good component $\Rightarrow$ **cross-store shopping**
  - **40%** due to store component $\Rightarrow$ **store-choice**
  - **10%** due to transaction component $\Rightarrow$ **purchase timing**
- They conclude that there seems to be

  "*significant variation in households' abilities to systematically take advantage of persistent price differences for the same good at different stores by purchasing each good at the store where that particular good is, on average, cheaper*"
- Intriguing result for two reasons:

How do people save? IV

- important role for cross-store shopping
- very small variation due to intertemporal optimization

Table of contents

1 Introduction

2 Our contribution

3 The method

4 Understanding the decomposition

5 An example

6 Concluding remarks

A finer decomposition I

- We argue that the KM decomposition is too coarse
- In particular, their store-good component conflates two different effects:
  - the ability to choose the right product from the right store (cross-store shopping)
  - the ability to choose the right store *for one's basket*
- Key insight: store expensiveness is not universal, but may differ across consumers depending on the basket they consume; in our data, 26% of consumer baskets cost less in a store that is more expensive according to a general price index
- Hence, the KM conclusion that there is significant variation in households' ability for cross-store shopping is tainted by the second effect

A finer decomposition II

- We propose a finer decomposition that breaks up the store-good component into these two parts:
  - **Store-basket component**: measures the expensiveness of the consumer's basket at the stores she visits relative to the average expensiveness of those stores
    $\Rightarrow$ **basket-based store choice**
  - **Pure store-good component**: measures the expensiveness of products purchased by the consumer relative to the store's basket-specific expensiveness
    $\Rightarrow$ **pure cross-store shopping**
- We applied both decompositions to the IRI Marketing Dataset (smaller than but similar to the Nielsen dataset used by KM)

A finer decomposition III

- Decomposition results (simplified version, without the transaction component):

	KM	CCM
Store choice	49	49
Cross-store shopping	55	24
Basket-based store-choice		46

- With the KM decomposition, the variance is explained by cross-store shopping and store choice in roughly equal amounts
- In the CCM decomposition, cross-store shopping (pure store-good component) explains a modest 24% of the variance, less than half what we find with KM.
- Basket-based store choice (our new store-basket component) comes in as a substantial contributor to variance at 46%.

A finer decomposition IV

- These findings suggest that variation in cross-store shopping is ***not*** as large as KM found.
- Rather, *households differ more in their ability to save by selecting stores that are cheap for the basket they purchase*.
- Comforting: consumers can obtain substantial savings without having to resort to product-level, cross-store price comparisons

Table of contents

1. Introduction
2. Our contribution
3. The method
4. Understanding the decomposition
5. An example
6. Concluding remarks

The KM decomposition

$$p_{i,t} = \underbrace{\sum_{j,k \in i \cap t} \mu_{j,m,t}\, \omega_{i,j,k,t}}_{\text{market average price of } j} + \underbrace{\sum_{j,k \in i \cap t} \left(\mu_{s,t} - \mu_{j,m,t}\right) \omega_{i,j,k,t}}_{\text{store}}$$
$$+ \underbrace{\sum_{j,k \in i \cap t} \left(\mu_{j,s,t} - \mu_{s,t}\right) \omega_{i,j,k,t}}_{\text{store-specific good}} + \underbrace{\sum_{j,k \in i \cap t} \left(p_{j,k} - \mu_{j,s,t}\right) \omega_{i,j,k,t}}_{\text{transaction}}, \tag{1}$$

where $\omega_{i,j,k,t}$ is the share of household i's expenditure that goes to product j on shopping trip k in period t

Hypothetical price indexes I

- We first rewrite the KM decomposition in a more intuitive way
- We do this by defining *hypothetical HPIs* that give us the cost of the consumer's basket under different shopping scenarios:

  p_i^m is the cost of the consumer's basket had she bought each item at the average market price, and is equal to unity by definition;

  p_i^{sg} is the cost of the consumer's basket had she bought each item at the average price of the store she purchased it from;

  p_i^s is the cost of the consumer's basket on the basis of the average expensiveness of the stores she purchases each item from; put differently, it is the average expensiveness of the stores the consumer visits, evaluated on the basis of her basket;

Hypothetical price indexes II

p_i^{sb} is the cost of the consumer's basket had she purchased all items in her basket in each of the stores she visits in proportion to her overall spending in each store; it measures the expensiveness of the panelist's basket at the stores visited.

- We can now rewrite the KM decomposition of the HPI in terms of the price indexes:

$$p_i = p_i^m + \underbrace{(p_i - p_i^{sg})}_{\text{transaction}} + \underbrace{(p_i^{sg} - p_i^s)}_{\text{KM store-good}} + \underbrace{(p_i^s - p_i^m)}_{\text{store}}. \tag{2}$$

- We simplify to focus on the store-good component:

$$p_i^{sg} - p_i^m = \underbrace{(p_i^{sg} - p_i^s)}_{\text{KM store-good}} + \underbrace{(p_i^s - p_i^m)}_{\text{store}}, \tag{3}$$

The CCM decomposition I

- We argue that, if our objective is to assess the household's ability to choose cheap products, then p_i^s is not the best benchmark to compare p_i^{sg} to.
- The reason is that p_i^s is based on a measure of store expensiveness, μ_s, that is calculated on the basis of *all* products.
- It is more appropriate to compare p_i^{sg} to our proposed new index p_i^{sb}, which is calculated using a *household-specific* measure of store expensiveness, $p_{i,s}$.
- Our innovation is to introduce p_i^{sb} and use it to define a finer decomposition of $p_i^{sg} - p_i^m$:

$$p_i^{sg} - p_i^m = \underbrace{(p_i^{sg} - p_i^{sb})}_{\text{pure store-good}} + \underbrace{(p_i^{sb} - p_i^s)}_{\text{store-basket}} + \underbrace{(p_i^s - p_i^m)}_{\text{store}} \tag{4}$$

The CCM decomposition II

The interpretation of each difference is as follows:

$p^s_{i,t} - p^m_{i,t}$ is the *store* component. It is the difference between the average cost of the stores visited by the panelist and cost of the average store in the market. It measures the panelist's ability to select cheap stores.

$p^{sg}_{i,t} - p^{sb}_{i,t}$ is the *pure store-good* component. It is the difference between the average cost of the goods the consumer purchases at the stores where each purchase is made and the cost of her basket at the stores she visits. It measures the panelist's ability to purchase each good in her basket at the store where that good is relatively cheaper within the set of stores visited.

$p^{sb}_{i,t} - p^s_{i,t}$ is the *store-basket* component. It is the difference between the average cost of the panelist's basket at the stores she visits and the average expensiveness of the stores she visits. It measures the extent to which the panelist purchases a basket that is representative of the expensiveness of the stores she visits.

Table of contents

1. Introduction
2. Our contribution
3. The method
4. Understanding the decomposition
5. An example
6. Concluding remarks

Understanding the decomposition I

- The decomposition is atheoretical and the economic intuition is hard to grasp
- Part of our contribution is to elucidate the inner workings of the methodology and show how it can inform the literature on consumer search
- Any search protocol will leave a footprint on the data in terms of the sizes of the components
- We outline simple models of consumer search (shopping protocols) and discuss how they would give rise to variation in different components

Understanding the decomposition II

Case	Prices	Consumers	Outcome
1	Same in all stores	All have same basket All randomize across stores	No variation in HPI
2	Different overall store price levels	All have same basket Shoppers choose the cheapest store Busy consumers randomize across stores	Variation in the store component
3	Different overall store price levels	All have different baskets All buy from a single store: Shoppers choose cheapest store for their basket Busy consumers choose a random store	Variation in the store or store-basket component
4	Relative price dispersion	All have same basket All visit same two stores: Shoppers buy each product where it is cheapest Busy consumers purchase same basket in both stores	Variation in the store-good component

Table of contents

1. Introduction
2. Our contribution
3. The method
4. Understanding the decomposition
5. An example
6. Concluding remarks

An illustrative example I

- Consider a market with two stores selling the same two products, say bread and milk.
- One store specializes in bread (call it a bakery) and the other specializes in milk (a dairy). Both stores offer lower prices on their specialty products.
- Consumers differ in two dimensions: the composition of their basket and their shopping behavior. One fifth of consumers are *shoppers* and the other four fifths are *loyals*.
- Loyals have a basket containing half a unit each of bread and milk. They buy from a single store that could be chosen, say, on the basis of location, and are evenly split across the two stores.

An illustrative example II

- Shoppers buy each item in their basket at the lowest available price. They come in three types of equal size: bread shoppers purchase one unit of bread from the bakery, milk shoppers one unit of milk from the dairy, and all-shoppers purchase a half-unit of milk from the dairy and a half-unit of bread from the bakery.

	(Stores)		Bakery		Dairy	
	(Products)		Bread	Milk	Bread	Milk
	Consumer type	*Frac.*	*Quantity purchased* q_{ijs}			
1	Bread shopper	1/15	1	0	0	0
2	Milk shopper	1/15	0	0	0	1
3	All-shopper	1/15	1/2	0	0	1/2
4	Dairy loyal	6/15	0	0	1/2	1/2
5	Bakery loyal	6/15	1/2	1/2	0	0
	Quantity purchased		3/10	1/5	1/5	3/10

An illustrative example III

- Suppose the price of bread at the bakery and milk ad the dairy is 1.0, while the price of bread at the dairy and milk at the bakery is 1.1.

	Price indexes					Components		
Cons. type	p	p^m	p^s	p^{sg}	p^{sb}	store-good KM $p^{sg} - p^s$	pure store-good CCM $p^{sg} - p^{sb}$	store-basket CCM $p^{sb} - p^s$
1	0.96	1.00	1.00	0.96	0.96	-0.04	0	-0.04
2	0.96	1.00	1.00	0.96	0.96	-0.04	0	-0.04
3	0.96	1.00	1.00	0.96	1.01	-0.04	-0.05	0.01
4	1.01	1.00	1.00	1.01	1.01	0.01	0	0.01
5	1.01	1.00	1.00	1.01	1.01	0.01	0	0.01
Market level decomposition						100%	39%	72%

An illustrative example IV

- Since the store component $p^s - p^m$ is zero by construction, the KM decomposition attributes price dispersion for all consumers entirely to the KM store-good component.
- This is appropriate for the all-shoppers (type 3) because these consumers save by purchasing the goods in their basket from the stores where these goods are cheap.
- But the nonzero KM store-good component for consumer types 1, 2, 4, and 5 is problematic. For example, consumers 1 and 2 have a negative KM store-good component because the cost of their single-item basket is lower than the overall expensiveness of the store they visit.
- Yet there is no basis on which to conclude that these consumers buy the right product from the right store, which is the KM interpretation of a nonzero store-good component, as they only purchase a single item from a single store.

An illustrative example V

- In contrast, the CCM pure store-good component is zero for all four consumer types (each type satisfies one of the conditions stated in Proposition 1).
- The subtle difference between the store-good and store-basket components leads to a striking difference between the two decompositions.
- The KM decomposition attributes 100 percent of the dispersion to the KM store-good component while the CCM decomposition attributes only 39 percent to the pure store-good component, with 72 percent being attributed to the store-basket component.
- The conclusion from the CCM decomposition is that differences in prices consumers pay is primarily due to variation in consumers' ability to select stores on the basis of the expensiveness of their basket in these stores, and less so to the ability of selecting the cheapest products across stores.

Models of cross-store shopping I

- Two models of consumer shopping:
  - Consumers compare prices of all products across all visited stores (Kaplan, Menzio, Rudanko, Trachter 2019 – KMRT)
  - Consumers shop across stores but at the product category level, e.g. food items from one store, cleaning supplies from another (Thomassen, Smith, Seiler, Schiraldi 2017 – TSSS)
- The two models have different implications for the decomposition applied separately to each product category:
  - Under KMRT: the decomposition should produce the same results when categories are examined together or separately
  - Under TSSS: the decomposition by category should have a zero pure store-good component.

Models of cross-store shopping II

- We applied the decomposition to five product categories:

	Carbonated soft drinks	Cereal	Milk	Salty snacks	Yogurt	Overall
Cross-store shopping	11	15	16	13	8	24
Basket-based store choice	43	53	69	47	64	46
Store choice	56	41	33	49	31	49

- The decrease in the role of the store-good component across the board is consistent with the TSSS search protocol, where consumers shop for categories rather than individual goods.
- But the fact that the pure store-good component is not zero suggests that some consumers do compare prices of the same good across stores, as in the KMRT model.

The full decomposition

- Results of full decomposition that includes the transaction component:

	KM	CCM
Store choice	20	20
Cross-store shopping	22	10
Purchase timing	51	51
Basket-based store-choice		19

- Again, cross-store shopping in CCM is less than half of what it is in KM
- Purchase timing large and similarly sized – bigger than with Nielsen data, remains a puzzle

Table of contents

1. Introduction
2. Our contribution
3. The method
4. Understanding the decomposition
5. An example
6. Concluding remarks

Summary

- Recent work has attributed a large fraction of variance in prices consumers pay in grocery stores to variation in cross-store shopping
- We claim that this finding is due to a failure to account for the possibility that consumers are choosing the right store for their basket
- Allowing for this possibility leads to lower variation due to cross-store shopping than previously found
- We also find evidence that much cross-store shopping is done at the category level, but some consumers also engage in cross-store shopping at the individual product level
- The idea that consumers can choose the cheapest store for their baskets seems powerful and merits further consideration.